



Disciplina: T951 - Sistemas Inteligentes

Professor: André Coelho

Datas de apresentação e entrega do código-fonte: 26/03/2019 e 28/03/2019

Peso da atividade: 4,0

Observação: Vale lembrar que **não** há segunda chamada para avaliações práticas. A **primeira data** será alocada para apresentação dos resultados relativos à tarefa de **regressão**, enquanto a **segunda data** para apresentação dos resultados relativos à tarefa de **classificação**. O **código-fonte** deverá ser entregue ao professor durante as aulas indicadas acima, juntamente com o **relatório técnico** reportando os resultados experimentais. **Todos** os alunos devem estar preparados para responder às perguntas feitas pelo professor, sendo que a **nota** de cada integrante de um grupo será dada com base em seu desempenho individual durante a arguição. **Não** será permitido o uso de *toolkits* ou bibliotecas de redes neurais disponíveis na Internet para implementar os algoritmos. O não-comparecimento implicará em nota **zero** nesta avaliação.

INSTRUÇÕES

Nesta atividade computacional, que poderá ser realizada individualmente ou em grupo de **no máximo 2 integrantes**, deve-se implementar em **Python, Java ou Matlab** os algoritmos de treinamento e teste das redes **Perceptron Multicamadas (PMC)** e **Adaline**, conforme o material disponibilizado em sala de aula. Os grupos devem se cadastrar até o dia **21/03/2019** via mensagem eletrônica para acoelho@unifor.br, indicando os nomes dos integrantes. Serão repassados dois *datasets*, um de regressão e outro de classificação, os quais serão descritos brevemente abaixo. Os resultados experimentais deverão ser apresentados de forma organizada, alocando uma seção diferente do relatório a cada experimento. **OBS:** o programa deve ser capaz de plotar a curva “EQM \times época”.

1. PROBLEMA DE REGRESSÃO – REDES ADALINE \times PERCEPTRON MULTICAMADAS

Nesta tarefa, será adotado o *dataset bodyfat* para avaliar comparativamente o potencial das redes Adaline e PMC em um problema de aproximação funcional (regressão múltipla). O problema é o de estimar a percentagem de gordura corporal de uma pessoa (variável dependente) a partir de 13 atributos físicos (variáveis independentes), os quais são descritos a seguir:

1. Idade (anos)	2. Peso (libras)	3. Altura (polegadas)	4. Circunferência do pescoço (cm)	5. Circunferência do peitoral (cm)
6. Circunferência do abdômen (cm)	7. Circunferência do quadril (cm)	8. Circunferência da coxa (cm)	9. Circunferência do joelho (cm)	10. Circunferência do tornozelo (cm)
11. Circunferência do bíceps estendido (cm)	12. Circunferência do antebraço (cm)	13. Circunferência do pulso (cm)		

No total, há 252 amostras disponíveis (não-normalizadas): 75% delas serão usadas no treinamento das redes e as demais para testar a capacidade de generalização. A partição treino/teste será definida pelo professor. Serão avaliadas 3 topologias da rede PMC, todas com uma camada escondida: PMC1, com 2 neurônios escondidos; PMC2, com 5 neurônios escondidos; e PMC3, com 10 neurônios escondidos. Deve-se usar a função tangente hiperbólica para os neurônios escondidos e linear para o neurônio de saída das redes PMC. O número máximo de épocas de treinamento deve ser 10000. A precisão a ser adotada para todas as redes deve ser $\varepsilon = 10^{-6}$. Em todas as execuções, os pesos devem ser iniciados com valores aleatórios em $[-0.5, 0.5]$. Em cada execução, reinicie o gerador de números aleatórios, de tal forma que as condições iniciais não sejam as mesmas. A partir dessas considerações, deve-se conduzir os experimentos descritos abaixo separadamente:

1) Considerando os dados normalizados via padronização e a taxa de aprendizado variando em $\eta = \{0.01, 0.1, 0.2, 0.5, 0.7, 1.0\}$, realize 5 execuções de treinamento (modo online) tanto para a rede Adaline como para cada topologia de rede PMC dada acima. Para as redes PMC, é necessário que as amostras de treinamento sejam embaralhadas a cada época. Em cada execução, anote o número de épocas de treinamento e

o valor do erro quadrático médio (EQM) final. Teste cada rede treinada sobre os dados de teste e também calcule o valor do EQM obtido sobre esse conjunto de dados. Não se esqueça que os dados de teste devem ser normalizados também. Finalmente, calcule os valores médios das medidas acima.

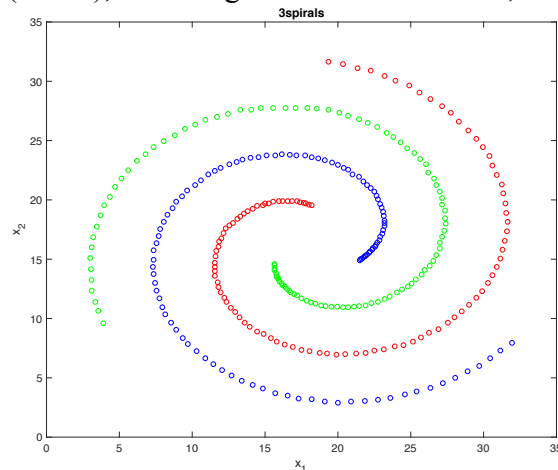
2) Repita o experimento anterior, considerando agora as redes treinadas no modo offline. Não é necessário embaralhar as amostras a cada época.

3) Repita os experimentos 1) e 2), somente para as redes PMC, considerando agora a normalização min-max no intervalo $[-0.5, 0.5]$.

4) A partir dos experimentos anteriores, indique qual topologia de rede PMC foi a mais eficaz, ou seja, obteve o menor erro de aproximação médio sobre as amostras de treinamento e teste. Houve alguma configuração de rede PMC (topologia + parâmetros de controle) cujo treinamento levou a *overfitting*?

2. PROBLEMA DE CLASSIFICAÇÃO – REDES PERCEPTRON MULTICAMADAS

Nesta tarefa, será adotado o *dataset 3spirals* para avaliar o potencial de diferentes topologias de redes PMC em um problema *benchmark* de classificação de padrões. No total, há 312 amostras disponíveis, sendo que 101 pertencem à primeira espiral (classe), 105 à segunda e 106 à terceira, conforme a figura a seguir:



A partição treino/teste dos dados será definida pelo professor, sendo 75% das amostras de treinamento e as demais de teste. Serão avaliadas 4 topologias da rede PMC: a topologia PMC1, com 6 neurônios de uma única camada escondida; a topologia PMC2, com 12 neurônios de uma única camada escondida; a topologia PMC3, com 6 neurônios na primeira camada escondida e 2 na segunda camada escondida; e a topologia PMC4, com 4 neurônios na primeira camada escondida e 3 na segunda camada escondida. Deve-se adotar codificação bipolar para as classes e a função tangente hiperbólica para todos os neurônios. O número máximo de épocas de treinamento deve ser 10000. A precisão a ser adotada deve ser $\varepsilon = 10^{-6}$. Em todas as execuções, os pesos devem ser iniciados com valores aleatórios em $[-0.5, 0.5]$. Em cada execução, reinicie o gerador de números aleatórios, de tal forma que as condições iniciais não sejam as mesmas. A partir dessas considerações, deve-se conduzir os experimentos descritos abaixo separadamente:

1) Considerando os dados normalizados via normalização min-max no intervalo $[-0.5, 0.5]$ e a taxa de aprendizado variando em $\eta = \{0.01, 0.1, 0.2, 0.5\}$, realize 5 execuções de treinamento (modo offline) para cada topologia de rede PMC descrita acima. Deve-se adotar a codificação sequencial das classes. Em cada execução, anote o número de épocas de treinamento e a taxa de classificação correta obtida ao final do treinamento. Teste cada rede treinada sobre os dados de teste e também calcule a taxa de classificação correta sobre esse conjunto de dados. Não se esqueça que os dados de teste devem ser normalizados também. Finalmente, calcule os valores médios das medidas acima.

2) Repita o experimento anterior, considerando agora a codificação "one-of-c" das classes.

3) Repita o experimento 2), adotando agora o termo de momentum na regra de ajuste dos pesos. A taxa de momentum deve variar em $\alpha = \{0.5, 0.7, 0.9\}$.

4) A partir dos experimentos anteriores, indique qual topologia de rede PMC foi em geral a mais eficaz, ou seja, obteve a maior taxa média de classificação correta sobre as amostras de treinamento e teste. Houve alguma configuração de rede PMC (topologia + parâmetros de controle) cujo treinamento levou a *overfitting*?